Signature properties of water

Signature properties of water: Their molecular electronic origins

Vlad P. Sokhan[a,1], Andrew P. Jones[b], Flaviu S. Cipcigan[b], Jason Crain[b], and Glenn J. Martyna[c]

[a]National Physical Laboratory, Teddington, Middlesex TW11 0LW, United Kingdom; [b]School of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom; and [c]IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598

Edited by Paul Madden, University of Oxford, Oxford, United Kingdom, and accepted by the Editorial Board March 31, 2015 (received for review October 1, 2014)

Water challenges our fundamental understanding of emergent materials properties from a molecular perspective. It exhibits a uniquely rich phenomenology including dramatic variations in behavior over the wide temperature range of the liquid into water's crystalline phases and amorphous states. We show that many-body responses arising from water's electronic structure are essential mechanisms harnessed by the molecule to encode for the distinguishing features of its condensed states. We treat the complete set of these many-body responses nonperturbatively within a coarse-grained electronic structure derived exclusively from single-molecule properties. Such a "strong coupling" approach generates interaction terms of all symmetries to all orders, thereby enabling unique transferability to diverse local environments such as those encountered along the coexistence curve. The symmetries of local motifs that can potentially emerge are not known a priori. Consequently, electronic responses unfiltered by artificial truncation are then required to embody the terms that tip the balance to the correct set of structures. Therefore, our fully responsive molecular model produces, ~~to our knowledge,~~ ~~the~~ first simple and intuitive picture of water's complexity and its molecular origin, predicting water's signature physical properties from ice, through liquid–vapor coexistence, to the critical point.

(2) (3) (1)
intermolecular interactions | many-body dispersion | subcritical water | coarse-grained model | electronic responses
(4)

Water is a ubiquitous yet unusual substance exhibiting anomalous physical properties for a liquid and forming many crystalline ices and (at least) two distinct amorphous states of different density (1). As the biological solvent, it is critical that water molecules form a liquid over a very wide range of temperatures (2) and pressures (3, 4) to support life under a wide variety of conditions. Indeed, water's simple molecular structure, a three-atom, two-species moiety, yields a surprisingly rich phenomenology in its condensed phases.

It is well-known that many signature properties of water have their molecular origin in the hydrogen-bonding interactions between molecules (5, 6). These directional networks are also the source of enhanced molecular polarization in the liquid state relative to the gas (7). In addition, there is speculation that dispersion interactions which arise from quantum-mechanical fluctuations of the charge density are also an important factor in the equilibrium properties of the ambient liquid (8, 9). The question of the ranges of temperature and density where these interactions influence observable properties is important for the construction of a conceptually simple but broadly transferable physical model linking molecular and condensed phase properties with the minimum of additional assumptions. Liquid water exhibits anomalies at both extremes of temperature—including a point of maximum density near freezing, an unusually high critical temperature relative to other hydrides, and significant changes in physical properties along the coexistence curve—thereby presenting a unique challenge for predictive simulation and modeling (10–13).

Today, simulation and modeling are considered the third pillar of the scientific method, together with analytical theory and experiment. Studies of condensed phase molecular systems via the combination of multiscale descriptions and statistical sampling have led to insights into physical phenomena across biology, chemistry, physics, materials science, and engineering (14). Significant progress now requires novel predictive models with reduced empirical input that are rich enough to embody the essential physics of emergent systems and yet simple enough to retain intuitive features (14).

Typically, atomistic models of materials are derived from a common strategy. Interactions are described via a fixed functional form with long-range terms taken from low orders of perturbation theory and are parameterized to fit the results of ab initio computations on test systems (both condensed and gas phase) and/or physical properties of systems of interest (15, 16). This approach presumes that the physics incorporated by the functional form is transferable outside the parameterization regime. In addition, first principles methods efficient enough to treat the condensed phase will also miss diagrams leading to truncation (e.g., local density functional theory neglects dispersion). If the truncation scheme is inappropriate for the problem of interest, predictions can go significantly awry (17, 18). Clearly, this is likely to be true for water with the directional, locally polarizing, H-bonded network of its liquid being disrupted in the less-associated gas phase. Consequently, the key physics underpinning a fully predictive model of water has yet to be identified.

The strategy adopted here captures water's properties from an atomistic perspective by incorporating individual water molecules, designed to respond with full many-body character, which can assemble to form condensed phases, as depicted in Fig. 1 *A*–*C*. This is a radical departure from standard approaches given above (11, 12). It is enabled by recent work (19–21) which has shown that it is possible to represent the complete hierarchy of long-range responses within a many-body system using a coarse-grained projection of the electronic structure onto an interacting set of

Significance

Water is one of the most common substances yet it exhibits anomalous properties important for sustaining life. It has been an enduring challenge to understand how a molecule of such apparent simplicity can encode for complex and unusual behavior across a wide range of pressures and temperatures. We reveal that embedding a complete hierarchy of electronic responses within the molecule allows water's phase behavior and signature properties to emerge naturally even within a simple model. The key result is, ~~to our knowledge, the first~~ prediction of liquid–gas phase equilibria from freezing to the critical point thus establishing a direct link between molecular and condensed phase properties and a sound physical basis for a conceptually simple but broadly transferable model for water.

Author contributions: V.P.S., ~~A.P.J.~~, J.C., and G.J.M. designed research; V.P.S. and F.S.C. performed research; V.P.S., A.P.J., F.S.C., J.C., and G.J.M. analyzed data; V.P.S., A.P.J., F.S.C., J.C., and G.J.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. P.M. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: vlad.sokhan@npl.co.uk.

quantum oscillators, one per single molecule—an approach that can be handled nonperturbatively and parameterized to the dilute gas limit using monomer molecular properties only. Thus, within Gaussian statistics, many-body responses arising from distortions of the electronic charge distribution (many-body polarization) and correlated quantum-mechanical charge density fluctuations (many-body dispersion) are included to all orders in the condensed phase, along with nontrivial cross-interactions—a novel strong coupling approach that has yet to be explored in molecular simulation. A description of the monomer electrostatics is added to capture lower-order molecular moments via point charges embedded in a rigid molecule frame, an approach borrowed from seminal early work (5, 10). A short-range pair potential determined from a single-dimer energy surface completes the quantum Drude oscillator (QDO) picture of water. All parameters used to simulate the model are given in Table 1—the parameters were not fit exhaustively through, for example, a machine-learning procedure which adds interest to the results presented below. Reasons for successes (or failure) and region of validity of the model are given following the presentation of the data.

Results

Properties Along the Liquid–Gas Coexistence Curve. Water's liquid–vapor coexistence curve in Fig. 2*A* has several unique features: The curve is wide and the critical point remarkably high for a molecular liquid; there is also a temperature of maximum density (TMD) shown in Fig. 2*A*, *Inset*, a characteristic anomaly of water. In Fig. 3 *A*–*C*, the local structure is given in terms of the partial radial distribution functions which indicate strong tetrahedral coordination. The dielectric constant varies markedly along coexistence (Fig. 4), reflecting significant changes in polarity and local order. The textbook explanation for these features is that as a hydrogen-bonding, strongly polar fluid, composed of electronically polarizable entities, water can be expected to form a highly stable liquid phase, with van der Waals interactions providing additional cohesion that increases in importance at higher temperatures; this picture certainly gives guidance but is not predictive. We shall show that these elements arise within the QDO model in such a way as to capture quantitatively the experimental results and, as we shall see, provide ~~new~~ insight.

The many-body features of the QDO model enable simultaneous determination of the coexistence densities along both branches of the binodal curve over the full range of coexistence

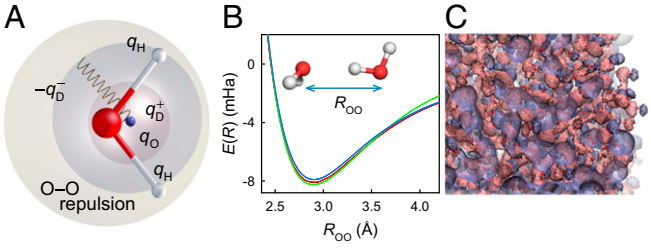

Fig. 1. QDO water model and predictions. (*A*) Schematic of the QDO water model, where a coarse-grained electronic structure, the Drude oscillator, is embedded in a rigid molecular frame decorated with point charges. The point charges capture the low-order electrostatic moments of the isolated molecule. The oxygen charge is placed on the *M* site down the symmetry axis, which is represented as a blue dot. The QDO is tethered to the *M* site. (*B*) The ground-state energy surface of the water dimer as a function of O–O distance with the molecular orientation fixed in that of the minimum energy geometry. QDO model results (red) are compared with high-level ab initio data (green) and polarizable Thole-type model potential TTM3-F, v. 3.0 (50) (blue). (*C*) Coarse-grained description of the changes in the electron distribution arising at the air–water interface; pink regions denote an increase and blue regions a decrease in electron density.

Table 1. QDO model for water principal parameters

Parameter	Value
R_{OH}	1.8088 a_0
∠HOH	104.52°
q_H	0.605 |e|
R_{OM}	0.504 a_0
m_D	0.3656 m_e
ω_D	0.6287 $E_h/\hbar$
$q_D^{\pm}$	±1.1973 |e|
κ_1	613.3 E_h
λ_1	2.3244 a_0^{-1}
κ_2	10.5693 E_h
λ_2	1.5145 a_0^{-1}
$\sigma_D = \sigma_H = \sigma_M$	0.1 a_0
σ_c	1.2 a_0

as given in Fig. 2*A*. We emphasize that the liquid–vapor density differences are very large near ambient conditions inducing extreme molecular-scale alterations, both in local spatial order (hydrogen bonding) and electronic structure. We determine the critical constants of the QDO model, $\{T_c, \rho_c\}$, by fitting the Wegner expansion (22) to our data, which express the difference between liquid and vapor densities (ρ_l and ρ_v, respectively), as $\rho_l - \rho_v = A_0\tau^{\beta} + A_1\tau^{\beta+\Delta}$, where $\tau = 1 - T/T_c$, $\beta \approx 0.325$ for the Ising universality class, and $\Delta = 0.5$. The results, $\{T_c = 649(2)$ K, $\rho_c = 0.317(5)$ g/cm$^3\}$, agree well with experimental values, $\{T_c = 647.096$ K, $\rho_c = 0.322$ g/cm$^3\}$ (23).

Because the baseline of the QDO model is the dilute limit (e.g., a single free molecule and an isolated dimer), excellent description of the low-temperature vapor densities is no surprise. However, the emergence of a highly realistic liquid and dense vapor branch is a model prediction because no liquid branch data of any sort were used in building the molecular model. Thus, the QDO approach can treat the strong variations in hydrogen bonding, electrostatic molecular moments (permanent plus induced), and van der Waals forces that underpin water's phenomenology.

Temperature of Maximum Density. In Fig. 2*A*, *Inset* we extend, at ambient pressure, the temperature range below ambient temperature to find that water's density maximum, which at ambient pressure exists at 277.13 K (23), also emerges naturally. The estimated TMD is 278.6(20) K with the uncertainty in the last digits shown in parentheses. The molecular origins of this phenomenon are a topic of current research with an observed correlation to tetrahedral coordination (24) which our model describes due to its geometry and the accurate gas-phase dipole and quadrupole moment (components) and their enhancement in the condensed phase. The density maximum is a further illustration of the subtle balance between competing forces (directional H bonding, polarization, and more isotropic dispersion interactions including their many-body forms) and is a key property that should be predicted by a representative model of water.

Structure and Thermodynamics. The QDO water structure under ambient conditions is compared with the most recently reported diffraction measurements (25, 26) in Fig. 3. The results are characteristic of a highly structured molecular liquid with a well-developed hydrogen-bond network. Additional forces arising from the more isotropic dispersion forces act to maintain an equilibrium density close to the experimental value. Thus, the inclusion of many-body interactions to all orders enables the model to describe both local structure and the phase diagram.

Along coexistence, water exhibits a marked temperature dependence in its relative static dielectric permittivity and vaporization enthalpy (Fig. 4). This dependence arises from strong coupling of

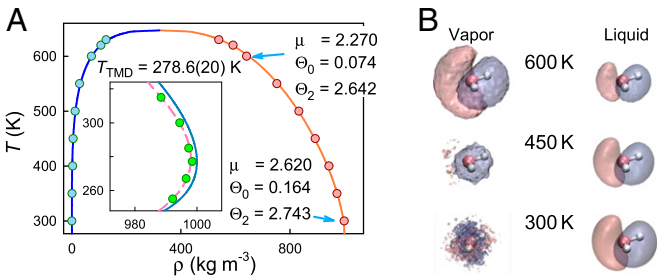

Fig. 2. Coexistence densities and the TMD of water. (*A*) Comparison of calculated orthobaric densities (symbols) with experimental data (solid lines). The error bars are smaller than the symbol size. The condensed phase molecular dipole and quadrupole moments at 300 and 600 K, as indicated by arrows, are also given. For reference, the model's gas phase values are $\mu = 1.855$ D, $\Theta_0 = -0.08$ DÅ, and $\Theta_2 = 2.49$ DÅ. (*Inset*) Density of water at ambient, $p = 0.1$ MPa, pressure in the temperature range around the experimental density maximum at 277.13 K; the full line is the compilation of experimental results (23, 51), the symbols are the QDO model results with one-sigma SD smaller than the symbol size. The TMD estimated from the spline fit (dashed line) is 278.6(20) K. (*B*) Change in electron density of water molecule in condensed phase along the vapor (*Left*) and liquid (*Right*) branches of the coexistence curve with respect to isolated molecule as a function of temperature. The statistically averaged snapshots are shown for $T = 300$, 450, and 600 K. The electron density gain is shown in pink, whereas density loss is given in blue.

water molecules to their local environment, depicted as the changes in the molecular electronic structure of Fig. 2*B* and in the tetrahedral coordination of Fig. 4, *Insets*. Obtaining the correct balance between the local structure and electronic structure changes over a wide temperature range 300 K $< T <$ 600 K is a prerequisite for prediction of these key properties.

The nontrivial electronic structure rearrangements that occur in water's condensed phase along coexistence are also reflected in its dielectric responses such as the permittivity given in Fig. 4 from experiment and the QDO model. Fig. 2*B* shows evolution of water's electronic structure with temperature from which induced electrostatic moments can be determined. The electrostatic moments are key to generating the correct dielectric behavior as the dielectric permittivity can be directly related to the fluctuations in the total dipole moment. The molecular dipole in the ambient liquid, for instance, is enhanced to $\mu = 2.6$ D from $\mu = 1.855$ D of the free molecule, a 40% increase. Both the dipole and quadrupole moment tensors in the ambient liquid predicted by the QDO model agree with recent high-level quantum-chemical estimates (27). Along the coexistence curve, there is an almost linear but modest depolarization of the water molecule whereby the dipole moment is reduced from $\mu = 2.6$ D to $\mu = 2.3$ D. The high-temperature moment is still considerably larger than that of the gas phase ($\mu = 1.855$ D), demonstrating that significant environmental correlations leading to molecular polarization are present, even in the disordered environment of the near-critical liquid (Fig. 2*B*) where proportionally more cohesion is provided by dispersion forces. The dipole moment under ambient conditions is consistent with other predictions (within 10%) (7, 28, 29), noting that the definition of a molecular dipole requires a model as an appropriate operator cannot be averaged over the many-body wavefunction to generate it.

We further find that the molecular properties which appear to be necessary to predict water's gas and vapor densities at a wide range of state points also generate realistic thermodynamics. This is evidenced by the temperature dependence along coexistence of the vaporization enthalpy, $\Delta H_{vap}(T)$, as shown in Fig. 4 in comparison with the reference (IAPWS) equation of state (23). At lower temperatures the hydrogen bonding and dipolar forces dominate, whereas at higher temperatures, van der Waals forces are a larger contribution. Capturing this fundamental force rebalancing demands an electronically responsive model.

Application to a Solid Polymorph: Proton-Ordered Ice II. Lastly, we apply the QDO approach in the context of a solid polymorph, the proton-ordered ice II—to avoid the complications arising from sampling disordered proton positions in this first application to a crystal (30). Ice II forms at pressures above ≈ 0.25 GPa and, when cooled below $T = 120$ K, it is metastable upon recovery to atmospheric pressure. Its structure is characterized by two distinct hexagonal rings of water molecules as shown in Fig. 5*A* together with a representation of the electronic rearrangement around the water molecules predicted by our coarse-grained electronic structure. The dimensions of the hexagonal unit cell of ice II are a and $b = a$, which control the size of the rings, and c, which controls the spacing between them. Fig. 5*B* shows the temperature dependence of both a and c, compared with the experimental data of Fortes et al. (31) over the temperature range 100 K $< T <$ 160 K. The hydrogen bonding, electrostatics, and many-body polarization roughly control a, whereas many-body dispersion allows c to emerge correctly.

Discussion

A simple model for water consisting of a rigid molecular frame decorated with point charges together with a coarse-grained electronic structure and a short-range pair potential, parameterized to a minimal set of monomer and dimer properties only, has been shown to predict the phase properties of liquid water over a wide range of thermodynamic states where the molecule remains, to good approximation, intact. The key insight is that such a minimal model can capture water's phase equilibria, dielectric properties along coexistence, its critical constants, a proton-ordered ice structure, and water's TMD. This suggests that the model contains the essential physics required to describe water over a wide variety of conditions. It remains to consider why the model is successful and to elaborate on the key physical principles this success reveals.

In assessing the QDO model's unique transferability, we begin with the basic principles underlying its construction. Embedding a QDO to treat long-range forces via a coarse-grained electronic structure and solving the QDO within a strong coupling approach (nonperturbatively) guarantees that interactions of all symmetry are present to all orders albeit within Gaussian statistics. Strong coupling solutions are important in physical models because they are unbiased by any a priori choice of dominant symmetry, order of interaction, or other such considerations. If the underlying model is sufficient, then predictive power across a variety of environments becomes possible. In general there are many configurations (free-energy basins) differing by minor free-energy differences in a complex system, and a strong coupling solution allows the dominant motifs for a given environment to emerge naturally as opposed to giving minor players artificial precedence due to

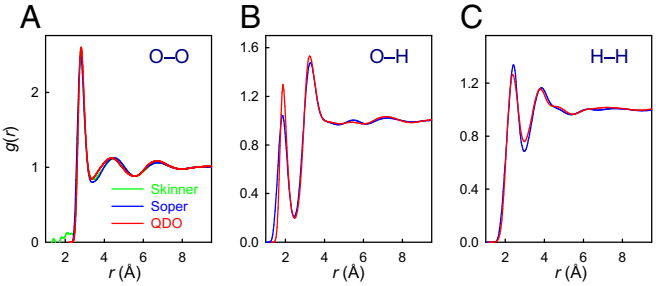

Fig. 3. (*A–C*) Partial radial distribution functions in the ambient QDO liquid (red lines) compared with recent neutron scattering (25) (blue lines) and X-ray diffraction (26) (green line) data. (*A*) Oxygen–oxygen; (*B*) oxygen–hydrogen, and (*C*) hydrogen–hydrogen.

IAPWS stands for The International Association for the Properties of Water and Steam (see http://www.iapws.org/)

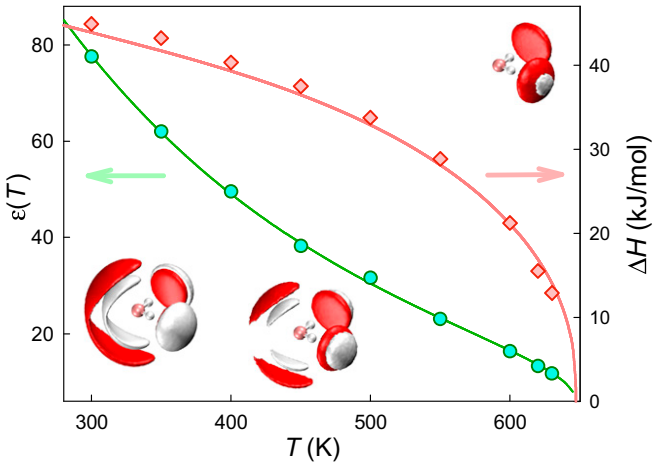

Fig. 4. Water properties along the coexistence line. The lines are the reference IAPWS equation of state (23); the symbols are the results of our simulations with the errors of the order of symbol size. Circles and left axis (indicated by arrows), relative static dielectric permittivity of water along the binodal; diamonds and right axis, enthalpy of vaporization of water. (*Insets*) Disruption of tetrahedral structure with increase of temperature along the binodal (T = 300, 450, and 600 K are shown) is illustrated.

truncation or symmetry selection imposed by the model builder and/or by a perturbative solution. In the absence of such a strong coupling approach, transferability becomes problematic because it is generally not known a priori which terms must be included explicitly.

To illustrate the value of a strong coupling approach, a representation of the evolution of the Drude particle charge distribution in the molecular frame along coexistence is presented in Fig. 2*B*. Within a dipole polarizable model (which is sometimes referred to as a "polarizable model" in the literature), the induced charge distribution can only exhibit dipole character by construction. The QDO model generates induced multipoles beyond dipole as can be clearly seen by the contours of the distributions and hence richer physics. Water's quadrupole moment tensor is in fact changed by the environment (by about 10% in principal components magnitude) and the QDO model is capable of capturing this effect (as well as all higher-order induced moments) within the approximation of Gaussian statistics (Fig. 2*A*).

The QDO model also captures many-body dispersion beyond the dipole approximation and to all orders (two-body, three-body, etc). These many-body terms, more important in the gas phase (high temperatures, low density) where the local environment is more spherical (21) and at surfaces where symmetry is broken, can tip the balance and yield improved model predictions although they themselves are not intrinsically large. Barker demonstrated (17) that it is insufficient to truncate the dispersion series at the pairwise level for noble gases as three-body terms are critical to phase equilibria and surface tension. Recently, the importance of many-body dispersion in aspirin polymorphism has also been revealed (18). By contrast, the radial distribution function in dense systems is not strongly affected, being driven by packing. These higher-order dispersion terms are less critical in ambient water where typical models truncate at the pairwise induced dipole-induced dipole level but, for the transferability across the phase diagram presented here, they matter. In other work (32), we have shown the QDO water model generates the surface tension along coexistence with high accuracy. This, along with all of the other physics (dielectric properties along coexistence and the surface tension predicted by the same model) is difficult to achieve outside a strong coupling approach.

The QDO water model works well without expending enormous effort in its parameterization because the QDO and the strong coupling solution used provide the long-range interactions (to all orders), thereby allowing the remainder (short-ranged repulsion) to be treated efficaciously within a pairwise approximation. Also, the large delocalized charge distribution of the quantum oscillator removes limitations arising from insufficient out-of-plane density in simple model electrostatics (27). Last, we have taken care to treat accurately the low-order moments (up to quadrupole) of the gas-phase charge distribution as their neglect has been identified as an important issue (27). It should be noted that the QDO model as currently constructed will fail (*i*) when the molecule dissociates (*ii*) at very high pressure, before dissociation, where three-body short-range repulsive terms will need to be added following, for example, the seminal work of Madden and co-workers (33). Rare events relying on molecular flexibility like certain dynamical processes around proteins (34) will also be outside of the predictive capability.

To provide further context for our findings, we note that early (and current) water models (5, 10) typically have fixed charges embedded in a rigid frame, fixed pairwise induced dipole–induced dipole dispersion terms (pairwise dispersion in the dipole approximation), and an oxygen-centered repulsion. Among the most important and challenging tests of model physics is the prediction of liquid–vapor equilibria as a function of temperature and of other physical properties along the coexistence curve. Here, the performance of this class of empirical model shows considerable variation. In general, the liquid–vapor coexistence curve for the most common water models (parameterized for the ambient liquid) severely underestimates the experimental density of the liquid branch with increasing temperature (35). The predictions suggest that either additional cohesive forces not contained in these models become more relevant under nonambient conditions or/and some aspects of the original model parameterization are inappropriate under circumstances where the H-bond network is reconfigured or weakened. More importantly, even the more successful models of this class do not immediately expose the relevant physics required to produce a better description of the balance of interactions governing liquid–vapor coexistence. The addition of polarizability in the dipole limit, for example, does not produce a substantial improvement. As a result, there has been no clear consensus as to the optimum parameterization strategy nor agreement on the extent to which it is reasonable to distort molecular properties of the isolated molecule to improve the description of the liquid under ambient conditions or other state points.

The above discussion suggests that a more complete representation of the electronic structure is required. However,

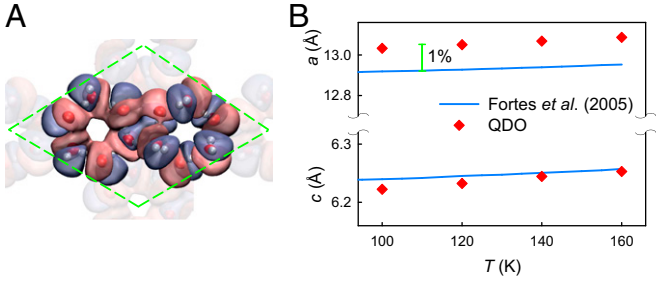

Fig. 5. Simulation of the proton-ordered ice II. (*A*) Characteristic hexagonal rings in ice II as viewed along the *c* axis of the hexagonal cell along with the coarse-grained electronic structure generated by our model. Regions of electron density increase are given in pink, whereas regions of depletion are given in blue. The distribution is defined as the difference between the total Drude oscillator density in the system and density of the unperturbed oscillators positioned on the molecular sites. (*B*) Unit cell dimensions of ice II crystals in the hexagonal settings as a function of temperature. The diamonds are the results of our simulations in the isobaric–isothermal ensemble at T = 100, 120, 140, and 160 K; the error bars are smaller than the symbols. The solid lines represent the experimental data obtained from ref. 31 for deuterated ice II.

current ab initio models sufficiently tractable for condensed phase applications neglect dispersion completely or include truncated approximations to the dispersion series. They do not, therefore, generate the balance of forces required to reproduce water's signature properties (36). On the other hand, a fully developed electronic structure incorporated at the level of simplicity afforded by a single embedded quantum oscillator, which can be solved nonperturbatively and efficiently, appears to be sufficient. Together with the model's fixed charge density (approximating the isolated molecule charge distribution), the QDO treatment generates the changing balance of textbook hydrogen-bonding, polar, and van der Waals forces required to account for the key properties of water across its vapor, liquid, and solid phases.

Given the results presented here and that our formulation is readily extensible to other materials and biological systems, we discuss the prospects for the QDO class of electronically coarse-grained models to yield ~~new~~ scientific insight into the emergence of complexity from the molecular perspective across the physical and life sciences. We highlight the following general considerations: (*i*) Models of the type we illustrate here are intuitive and transparent, having properties defined entirely in terms of isolated molecules—unbiased toward any thermodynamic state point or condensed phase. (*ii*) The resulting interactions are rich, containing the complete hierarchy of many-body inductive and dispersive forces as they are solved in strong coupling. (*iii*) They are capable of properly generating the balance between hydrogen-bonding, electrostatic, and van der Waals interactions; this, together with point (*i*), offers a promising foundation for a highly versatile description of matter, linking molecular physics to material properties. (*iv*) The simplification of the electronic problem inherent in the QDO model permits linear scale sampling methods to be used (37). The formulation can thus be applied with no loss of accuracy to large systems, allowing the molecular-scale exploration of a wide variety of important scientific phenomena at greater length scales of organization including hydrophobic hydration and drying (38), ionic solvation as well as Hofmeister effects, and biological processes such as protein–protein association.

Materials and Methods

The Model and the Simulation Method. The QDO-based multiscale Hamiltonian for a system with N water molecules is

$$H = \sum_i T_i^{(\mathrm{rigid})} + \phi^{(\mathrm{Coul})}(\mathbf{R}) + \sum_{j>i} \phi^{(\mathrm{rep})}(R_{Oij}) + E_0^{(\mathrm{Drude})}(\mathbf{R}),$$

$$\hat{H}^{(\mathrm{Drude})}\psi_0(\mathbf{r},\mathbf{R}) = E_0^{(\mathrm{Drude})}(\mathbf{R})\ \psi_0(\mathbf{r},\mathbf{R}), \qquad [1]$$

$$\hat{H}^{(\mathrm{Drude})} = \sum_i \left(\hat{T}_i + \frac{m_D \omega_D^2}{2}(\mathbf{r}_i - \mathbf{R}_{ci})^2 \right) + \phi^{(\mathrm{Coul})}(\mathbf{r},\mathbf{R}).$$

Here, $T_i^{(\mathrm{rigid})}$ is the classical rigid-body kinetic energy of water molecule i; Q:13 $9N$-vector $\mathbf{R}$ represents the coordinates of all of the water molecules in the system; $\phi^{(\mathrm{Coul})}(\mathbf{R})$ is the intermolecular Coulomb interaction energy between the fixed monomer gas-phase charge distributions only; $\phi^{(\mathrm{rep})}(R_{Oij})$ is an oxygen-centered pairwise repulsion. The quantities $E_0^{(\mathrm{Drude})}(\mathbf{R})$ and $\psi_0(\mathbf{r},\mathbf{R})$ are the ground-state Born–Oppenheimer energy surface and wavefunction of the QDO electronic structure, respectively; $\mathbf{r}_i$ and $\hat{\mathbf{T}}_i$ are the position and quantum kinetic energy of the oscillator centered on molecule i. The oscillators centered at $\mathbf{R}_{ci}$ are characterized by parameters $\{m_D, \omega_D, q_D\}$. Finally, $\phi^{(\mathrm{Coul})}(\mathbf{r},\mathbf{R})$ is intermolecular Coulomb interaction between the Drude particles of a negative charge of $-q_D$, at position $\mathbf{r}$, their centers of oscillation with a positive charge of $+q_D$, at position $\mathbf{R}_c$, and the fixed monomer charge distribution. Regularization of the Coulomb interaction is discussed elsewhere (20). All model parameters are given in Table 1.

We have treated the water molecules in our multiscale Hamiltonian as classical rigid bodies following seminal early work (10). This is a good approximation because the essential physics underlying water's properties is not altered in going from light water, H_2 O, to tritiated water, T_2 O. The critical constants, for instance, are not strong functions of isotope mass, T_c shifting 5 ° C and the temperature of maximum density by 9 ° C from H_2 O to T_2 O (1)—both the molecular polarizability and dispersion coefficients decrease with increasing isotope mass, counterbalancing quantum effects. In addition, it is known that there is a cancellation of effects as one goes from a classical rigid-body description of the water molecule to a quantum-mechanical, fully flexible molecular entity (36), hence the wide adoption of the classical rigid-body approach (10). The molecular vibrational degrees of freedom being close to ground state have little entropy as in the rigid model, and classical flexible, classical rigid, and quantum flexible models simulated under the same simple force law show the rigid approximation closely matches the quantum result, whereas the classical flexible model exhibits strong deviations (36, 39).

Classical Drude models are limited to dipole polarization and there are many such models in the literature (40–42) which can provide comparisons for our results. This model class, dipole polarizable models, is more computationally efficient than its quantum analog, but the responses are much more limited. We have derived exact expressions for dispersion coefficients and higher-order polarizabilities in the QDO model— all of which depend on Plank's constant and hence vanish in the classical limit (20). This provides a rigorous justification for the first sentence in this paragraph. The classical limit of the QDO, the classical Drude model, would have no dispersion and could not reproduce the dimer curve of Fig. 1 (without the ad hoc addition of pairwise dispersion, quadrupolar polarization, and higher terms).

The Drude oscillator parameters, $\{m_D, \omega_D, q_D\}$, are chosen such that the long-range responses match reference values of the monomer and dimer, i.e., monomer polarizabilities and pair-dispersion coefficients. The values of three fixed point charges, $\{q_H, q_H, q_O = -2q_H\}$, and the position of center of negative charge placed down the molecular bisector at position $\mathbf{R}_M$, are selected to represent the isolated molecule electrostatics by matching to low-order electrostatic moments; the positive charges are fixed on the positions of the hydrogen atoms. A QDO is tethered to the point $\mathbf{R}_M$ which coincides with the position of M site of the TIP4P model (21). The values of Q:14 the Gaussian charge widths are given in Table 1.

Short-range repulsion is incorporated as the difference between a reference high-level quantum-chemical dimer potential energy surface [calculated Q:15 at the CCSD(T) level using aug-cc-pVTZ basis set using ACESIII, Version 3.0.7 (43)], and the QDO dimer ground energy surface computed using norm-conserving diffusion Monte Carlo for the QDO model (44) with 1,000 walkers. The difference is represented as the isotropic pairwise, oxygen-centered term $\phi^{(\mathrm{rep})}(R_{Oij})$ of Eq. **1** and is approximated in the simulation by (44) $\phi^{(\mathrm{rep})}(r) = \sum_i \kappa_i \exp(-\lambda_i r)$, with parameters given in Table 1.

Molecular Dynamics Implementation with Adiabatic Path Integrals. Finite-temperature condensed phase simulation of the QDO model in the *NVT*, *NpT*, and *NpT*–flex ensembles were performed using the adiabatic path integral molecular dynamics for QDOs (APIMD-QDO) method (19, 37, 45). The technique uses a separation in time scale between the path integral degrees of freedom representing the Drude oscillators and the molecules, to generate motion of the water molecules on the coarse-grained electronic Born–Oppenheimer surface provided by the Drude oscillators. APIMD-QDO scales as $\mathcal{O}(N)$ in the number of molecules. The path integral was discretized using $P = 96$ beads. We typically studied a periodic system of $N = 300$ water molecules and used 3D-Ewald summation (46) with vacuum boundary conditions (47) (Ewald parameter $\epsilon_{\mathrm{rf}} = 1$) to compute the electrostatic interactions.

Simulation of Coexistence Densities and Thermodynamic and Dielectric Properties. To simulate water along the binodal, the coexistence pressure (or density) is required. We used direct simulation of two coexisting phases in a series of *NVT* calculations of water forming a slab in the central part of an elongated $L_z > \forall\{L_x, L_y\}$ cell (48), as a Gibbs ensemble method for QDOs has yet to be developed. The coexistence pressures obtained were then used in an independent series of *NpT* simulations of the bulk phase. Resulting densities were then compared with those obtained from hyperbolic tangent fits to the density profile of the slab (48), to assess consistency. The normal component of the stress tensor, calculated using the virial route (49), agrees, within statistical uncertainty, with experimental values. In all cases, the QDO water system was evolved for a simulation time of at least 1 ns.

The dielectric permittivity, ε, of bulk water was estimated using the linear response theory expression for a system with electrostatics computed using Ewald summation (46, 47),

$$\frac{(\epsilon - \epsilon_\infty)(2\epsilon_{\mathrm{rf}} + 1)^2}{(2\epsilon_{\mathrm{rf}} + \epsilon)(2\epsilon_{\mathrm{rf}} + \epsilon_\infty)} = \frac{\langle \overline{\mathbf{M}}_n^2 \rangle - \langle \overline{\mathbf{M}}_n \rangle^2}{3\epsilon_0 V k_B T}. \qquad [2]$$

Here, $\overline{\mathbf{M}}_n$ is the total dipole moment of the sample for a given set of molecular positions (see below), ϵ_{rf} is the permittivity of the medium at the asymptotic surface (36, 47), $V = L_x L_y L_z$ is the system volume, k_B is the Boltzmann constant, T is the thermostat temperature, ϵ_0 is the electric

constant, and the angle brackets represent the canonical ensemble average of the molecules on the ground-state energy surface provided by the quantum Drude oscillators at temperature T (performed statistically over at least 1 ns of simulation time). We evaluated the high-frequency limit ϵ_∞ using equation 24 of ref. 47 appropriately modified for our model. In the above calculations of the dielectric permittivity, performed in the *NVT* ensemble under 3D periodic boundary conditions, we also used $\epsilon_{rf} = 1$, taking the densities and temperatures from the coexistence simulations.

Within the APIMD-QDO method, we do not have access to the QDO properties for each instantaneous nuclear configuration; to obtain this information the path integral degrees of freedom would have to be formally integrated (or averaged) out. To perform the required average to good approximation on the total dipole moment operator of Eq. **2**, the staging dipole moment estimator of ref. 19 is subaveraged over a time interval, $\tau = n\delta t$, where $\delta t = 0.125$ fs is the APIMD-QDO time step and the subaverage is indicated by the notation $\overline{\mathbf{M}}$ in Eq. **2**. As the adiabatic separation (how fast the Drude path integral degrees of freedom evolve compared with the physical atoms) is selected such that the time scale of nuclear motion encompasses many correlation times of the path integral sampling, the approach converges rapidly with n, allowing $n = 20$ corresponding to $\tau = 2.5$ fs to be used; the averaging time scale is thus small compared with the Debye relation time of water (*ca*. 10 ps).

The enthalpy of vaporization, ΔH_{vap}, was calculated in the *NpT* ensemble using the densities and temperatures from the coexistence simulation. Because we assumed rigid monomer geometry in the QDO model, a temperature-dependent correction term for intra- and intermolecular vibrations (16) has been added.

Simulation of Ice II. To calculate the unit cell dimensions and density of ice II, we simulated a supercell of $N = 324$ water molecules at ambient pressure ($p = 0.1$ MPa) and temperatures ranging from $T = 100$ to 160 K, in steps of 20 K, using periodic boundary conditions and a fully flexible unit cell (*NpT*-flex ensemble). The supercell was equilibrated for a simulation time of 20 ps and the values reported in Fig. 3*B* are the result of averaging over a further 100 ps of simulation time. We used vacuum boundary conditions in Ewald summation of electrostatic forces ($\epsilon_{rf} = 1$), which resulted in minimal finite size effects.

ACKNOWLEDGMENTS. G.J.M. thanks Prof. Toshiko Ichiye and Dr. Dennis M. Newns for interesting discussions and the University of Edinburgh for an honorary Professorship in Physics. This work was supported by National Physical Laboratory Strategic Research Programme. A.P.J. acknowledges the European Metrology Research Programme support. We acknowledge use of Hartree Centre IBM BlueGene/Q resources in this work.

1. Franks F, ed (1972–1982) *Water: A Comprehensive Treatise* (Plenum, New York), Vol 1–7.
2. Brock TD (1967) Life at high temperatures. Evolutionary, ecological, and biochemical significance of organisms living in hot springs is discussed. *Science* 158(3804):1012–1019.
3. Rothschild LJ, Mancinelli RL (2001) Life in extreme environments. *Nature* 409(6823): 1092–1101.
4. Sharma A, et al. (2002) Microbial activity at gigapascal pressures. *Science* 295(5559): 1514–1516.
5. Bernal JD, Fowler RH (1933) A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J Chem Phys* 1:515–548. 8
6. Pauling L (1970) *General Chemistry* (Freeman, San Francisco), 3rd ed.
7. Gregory JK, Clary DC, Liu K, Brown MG, Saykally RJ (1997) The water dipole moment in water clusters. *Science* 275(5301):814–817.
8. DiStasio RA, Jr, von Lilienfeld OA, Tkatchenko A (2012) Collective many-body van der Waals interactions in molecular systems. *Proc Natl Acad Sci USA* 109(37):14791–14795.
9. DiStasio RA, Jr, Gobre VV, Tkatchenko A (2014) Many-body van der Waals interactions in molecules and condensed matter. *J Phys Condens Matter* 26(21):213202.
10. Stillinger FH, Rahman A (1974) Improved simulation of liquid water by molecular dynamics. *J Chem Phys* 60:1545–1557. 4
11. Errington JR, Debenedetti PG (2001) Relationship between structural order and the anomalies of liquid water. *Nature* 409(6818):318–321.
12. Bukowski R, Szalewicz K, Groenenboom GC, van der Avoird A (2007) Predictions of the properties of water from first principles. *Science* 315(5816):1249–1252.
13. Stone AJ (2007) Chemistry. Water from first principles. *Science* 315(5816):1228–1229.
14. Mahoney MW, Jorgensen WL (2001) Quantum, intramolecular flexibility, and polarizability effects on the reproduction of the density anomaly of liquid water by simple potential functions. *J Chem Phys* 115:10758–10768. 23
15. Jorgensen WL, Tirado-Rives J (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc Natl Acad Sci USA* 102(19):6665–6670.
16. Horn HW, et al. (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* 120(20):9665–9678.
17. Barker JA (1986) Many-body interactions in rare gases: Krypton and xenon. *Phys Rev Lett* 57(2):230–233.
18. Reilly AM, Tkatchenko A (2014) Role of dispersion interactions in the polymorphism and entropic stabilization of the aspirin crystal. *Phys Rev Lett* 113(5):055701.
19. Whitfield TW, Martyna GJ (2007) Low variance energy estimators for systems of quantum Drude oscillators: Treating harmonic path integrals with large separations of time scales. *J Chem Phys* 126(7):074104.
20. Jones AP, Crain J, Sokhan VP, Whitfield TW, Martyna GJ (2013) Quantum Drude oscillator model of atoms and molecules: Many-body polarization and dispersion interactions for atomistic simulation. *Phys Rev B* 87:144103. 14
21. Jones A, Cipcigan F, Sokhan VP, Crain J, Martyna GJ (2013) Electronically coarse-grained model for water. *Phys Rev Lett* 110(22):227801.
22. Wegner FJ (1972) Corrections to scaling laws. *Phys Rev B* 5:4529–4536.
23. Wagner W, Pruß A (2002) The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *J Phys Chem Ref Data* 31:387–535. 2
24. Mahoney MW, Jorgensen WL (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 112:8910–8922. 23
25. Soper AK (2013) The radial distribution functions of water as derived from radiation total scattering experiments: Is there anything we can say for sure? *ISRN Phys Chem* 2013:1–67.
26. Skinner LB, et al. (2013) Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J Chem Phys* 138(7):074506.
27. Niu S, Tan ML, Ichiye T (2011) The large quadrupole of water molecules. *J Chem Phys* 134(13):134501.
28. Onsager L, Dupuis M (1960) in Termodinamica dei processi irreversibili, *Proceedings of the International School of Physics "Enrico Fermi"* (N. Zanichelli, Bologna), Vol. 10, pp 294–315.
29. Badyal YS, et al. (2000) Electron distribution in water. *J Chem Phys* 112:9206–9208. Q:19 21
30. Rick SW, Haymet ADJ (2003) Dielectric constant and proton order and disorder in ice Ih: Monte Carlo computer simulations. *J Chem Phys* 118:9291–9296. 20
31. Fortes AD, Wood IG, Alfredsson M, Vočadlo L, Knight KS (2005) The incompressibility and thermal expansivity of D_2O ice II determined by powder neutron diffraction. *J Appl Cryst* 38:612–618. 4 Q:20
32. Cipcigan FS, Sokhan VP, Jones AP, Crain J, Martyna GJ (2015) Hydrogen bonding and molecular orientation at the liquid-vapour interface of water. *Phys Chem Chem Phys* 17(14):8660–8669.
33. Aguado A, Bernasconi L, Madden PA (2002) A transferable interatomic potential for MgO from ab initio molecular dynamics. *Chem Phys Lett* 356:437–444. 5
34. Ceriotti M, Cuny J, Parrinello M, Manolopoulos DE (2013) Nuclear quantum effects and hydrogen bond fluctuations in water. *Proc Natl Acad Sci USA* 110(39):15591–15596.
35. Vega C, Abascal JLF, Nezbeda I (2006) Vapor-liquid equilibria from the triple point up to the critical point for the new generation of TIP4P-like models: TIP4P/Ew, TIP4P/2005, and TIP4P/ice. *J Chem Phys* 125(3):34503.
36. Mantz YA, Chen B, Martyna GJ (2006) Structural correlations and motifs in liquid water at selected temperatures: Ab initio and empirical model predictions. *J Phys Chem B* 110(8):3540–3554.
37. Jones A, et al. (2013) Electronically coarse-grained molecular dynamics using quantum Drude oscillators. *Mol Phys* 111:3465–3477. 22--23 *(see footnote)
38. Yu H, van Gunsteren WF (2004) Charge-on-spring polarizable water models revisited: From water clusters to liquid water to ice. *J Chem Phys* 121(19):9549–9564.
39. Babin V, Leforestier C, Paesani F (2013) Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, VRT spectrum, and second virial coefficient. *J Chem Theory Comput* 9:5395–5403. 12
40. Sprik M, Klein ML (1988) A polarizable model for water using distributed charge sites. *J Chem Phys* 89:7556–7560. 12
41. Lamoureux G, Alexander D, MacKerell J, Roux B (2003) A simple polarizable model of water based on classical Drude oscillators. *J Chem Phys* 119:5185–5197. 10
42. Lamoureux G, Roux B (2003) Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J Chem Phys* 119: 3025–3039. 6
43. Lotrich V, et al. (2008) Parallel implementation of electronic structure energy, gradient, and Hessian calculations. *J Chem Phys* 128(19):194104.
44. Jones A, Thompson A, Crain J, Müser MH, Martyna GJ (2009) Norm-conserving diffusion Monte Carlo method and diagrammatic expansion of interacting Drude oscillators: Application to solid xenon. *Phys Rev B* 79:144119. 14 Q:21
45. Martyna GJ, Tuckerman ME, Tobias DJ, Klein ML (1996) Explicit reversible integrators for extended systems dynamics. *Mol Phys* 87:1117–1157. 5
46. de Leeuw SW, Perram JW, Smith ER (1980) Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc R Soc Lond A* 373:27–56. 1752
47. Neumann M, Steinhauser O (1984) Computer simulation and the dielectric constant of polarizable polar systems. *Chem Phys Lett* 106:563–569. 6
48. Chapela GA, Saville G, Thompson SM, Rowlinson JS (1977) Computer simulation of a gas-liquid surface. Part 1. *J Chem Soc, Faraday Trans II* 73:1133–1144. 7
49. Rowlinson JS, Widom B (1982) *Molecular Theory of Capillarity* (Clarendon, Oxford).
50. Fanourgakis GS, Xantheas SS (2008) Development of transferable interaction potentials for water. V. Extension of the flexible, polarizable, Thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *J Chem Phys* 128(7):074506.
51. Hare DE, Sorensen CM (1987) The density of supercooled water. II. Bulk samples cooled to the homogeneous nucleation limit. *J Chem Phys* 87:4840–4845. 8

* special double issue that combines 22 and 23 (see DOI: 10.1080/00268976.2013.843032)

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 1

Q: 1_Please contact PNAS_Specialist.djs@sheridan.com if you have questions about the editorial changes, this list of queries, or the figures in your article. Please include your manuscript number in the subject line of all email correspondence; your manuscript number is 201418982. Thank you

Q: 2_Please (i) review the author affiliation and footnote symbols carefully, (ii) check the order of the author names, and (iii) check the spelling of all author names, initials, and affiliations. Please check with your coauthors about how they want their names and affiliations to appear. To confirm that the author and affiliation lines are correct, add the comment "OK" next to the author line. This is your final opportunity to correct any errors prior to publication. Misspelled names or missing initials will affect an author's searchability. Once a manuscript publishes online, any corrections (if approved) will require publishing an erratum; there is a processing fee for approved erratum. Changes indicated

Q: 3_Please review and confirm your approval of the short title: QDO water. If you wish to make further changes, please adhere to the 50-character limit. (NOTE: The short title is used only for the mobile app and the RSS feed.) A quantum Drude oscillator model for water

Q: 4_Please review the information in the author contribution footnote carefully. Please make sure that the information is correct and that the correct author initials are listed. Note that the order of author initials matches the order of the author line per journal style. You may add contributions to the list in the footnote; however, funding should not be an author's only contribution to the work. Change indicated.

Q: 5_Reminder: You have chosen not to pay an additional $1350 (or $1000 if your institution has a site license) for the PNAS Open Access option. Please confirm this is correct and note your approval in the margin. OK.

Q: 6_PNAS allows up to five keywords. You may add 1 keyword. Also, please check the order of your keywords and approve or reorder them as necessary. keywords reordered and one new one added.

Q: 7_Please confirm whether all units, divisions, departments, laboratories, or sections have been included in the affiliations line for each footnote symbol or add if missing. PNAS requires smallest institutional unit(s) to be listed for each author in each affiliation. Changes are indicated.

Q: 8_Please note that italic font cannot be used for emphasis in text. Ok. It wasn't italic, it was \emph{} in LaTeX. PNAS LaTeX style converted it to italic.

Q: 9_PNAS discourages claims of priority; is this truly new? If not, please either (a) replace the term "new" with a term such as "previously unidentified" or (b) remove it altogether to avoid the claim of priority. Manuscript modified as indicated to comply

Q: 10_The citation numbers "28, 7" are out of order. OK.

Q: 11_PNAS articles should be accessible to a broad scientific audience. As such, please spell out IAPWS. see text

Q: 12_In sentence beginning "By contrast," is insertion of comma between "affected" and "being driven" correct? Yes

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 2

Q: 13_Should the hyphen between "9N" and "vector" be a minus sign? It is a hyphen

Q: 14_PNAS articles should be accessible to a broad scientific audience. As such, please spell out TIP4P. see text

Q: 15_PNAS articles should be accessible to a broad scientific audience. As such, please spell out CCSD (T), aug-cc-pVTZ, and ACESIII (if this is an abbreviation). This is standard nomenclature and would add no value to the paper (see below after full lines).

Q: 16_Please supply issue numbers for ref. 10 and all references from journals that assign issue numbers. Done in red. see text

Q: 17_Is ref. 20 a one-page article, a letter, or a published abstract? If it is an unpublished abstract, then it must be removed from the reference list and placed in a footnote to the text. Please indicate which type of article it is, and it will be fixed accordingly in the proofs. it is a paper - see text. Physical Review uses article numbers instead of pages Ref. 20 is on 18 pages, from 144103-1 to 144103-18

Q: 18_Please verify page range for ref. 23. Correct.

Q: 19_Please verify that all et al. references contain 6 or more authors. If 5 authors or fewer, please supply complete author lists. ok. see text. All references were done using BibTeX and PNAS BibTeX style, the "et al." were inserted by the PNAS-provided script.

Q: 20_Please note that original ref. 32 was a duplicate of ref. 17. The duplicate has been removed and all remaining references have been renumbered. No, they were different. Ref (32) DOI: 10.1103/PhysRevLett.57.230 Ref. (17) DOI: 10.1080/00268978600100541 - (Molecular Physics)

Q: 21_Is ref. 45 a one-page article, a letter, or a published abstract? If it is an unpublished abstract, then it must be removed from the reference list and placed in a footnote to the text. Please indicate which type of article it is, and it will be fixed accordingly in the proofs. It is a paper; see text

---

Q: 15

CCSD(T) = Coupled Cluster singles double with triples in perturbation theory.

ACES = Advanced Concepts in Electronic Structure.

aug-cc-pVTZ = augmented correlation consistent polarized valence triple zeta basis set.